CONF-780245--1

LA-UR-78-721

TITLE: ON CHERNOFF FACES

AUTHOR(S): Lawrence A. Bruckner

SUBMITTED TO: Symposium on Graphical Representation of Multivariate Data, Naval Postgraduate School, Monterey, California.

MASTER

los alamos
scientific laboratory
of the University of California
LOS ALAMOS, NEW MEXICO 87545

Form No. 836
St. No. 2629
1/75

UNITED STATES
ENERGY RESEARCH AND
DEVELOPMENT ADMINISTRATION
CONTRACT W-7405-ENG. 36

ON CHERNOFF FACES

Lawrence A. Bruckner*

University of California
Los Alamos Scientific Laboratory
Los Alamos, New Mexico

Herman Chernoff introduced the idea of using faces to represent multidimensional data in 1971. Since then, this technique has been used in a wide variety of applications.

The first part of this paper discusses how to use the technique. Then Andrews' sine curves and Anderson's metroglyphys are introduced and compared to the facial representations. Dependencies among the facial features are considered next and a way to eliminate dependencies presented. Finally, some uses of Chernoff Faces at the Los Alamos Scientific Laboratory are mentioned.

I. INTRODUCTION

The use of Chernoff faces to investigate multidimensional data has been accelerating in the last few years. The most common usages of the technique are to display the data in a convenient form, to aid in discovering clusters and outliers, and to show changes with time.

The idea of using faces to represent multidimensional data was introduced by Professor Herman Chernoff under a contract with the Office of Naval Research while at Stanford University in 1971 (1). Professor Chernoff considered data having a maximum of 18-dimensions and allowed each dimension to be represented by one of 18 facial features. A typical Chernoff face is presented in Fig. 1. Herbert T. Davis, Jr., added nose width and ears to the face while at the Los Alamos Scientific Laboratory (LASL) in 1975. This revised face is shown in Fig. 2.

*This work was supported by the Conservation Division of the U. S. Geological Survey, Denver, Co.

Table I identifies the facial features, the range of values each may assume, and the value the program assumes for the facial feature when that feature is not used to represent a data dimension.

TABLE I. Description of Facial Features and Ranges

Variable		Facial Feature	Default Value	Range	
x_1	controls h^*	face width	.60	.20	.70
x_2	controls θ^*	ear level	.50	.35	.65
x_3	controls h	half-face height	.50	.50	1.00
x_4	is	eccentricity of upper ellipse of face	.50	.50	1.00
x_5	is	eccentricity of lower ellipse of face	1.00	.50	1.00
x_6	controls	length of nose	.25	.15	.40
x_7	controls p_m	position of center of mouth	.50	.20	.40
x_8	controls	curvature of mouth	0.00	4.00	4.00
x_9	controls	length of mouth	.50	.30	1.00
x_{10}	controls y_e	height of center of eyes	.10	0.00	.30
x_{11}	controls x_e	separation of of eyes	.70	.30	.80
x_{12}	controls θ	slant of eyes	.50	.20	.60
x_{13}	is	eccentricity of eyes	.60	.40	.80
x_{14}	controls L_e	half-length of eye	.50	.20	1.00
x_{15}	controls	position of pupils	.50	.20	.80
x_{16}	controls y_b	height of eyebrow	.80	.60	1.00
x_{17}	controls $\theta^{**}-\theta$	angle of brow	.50	.00	1.00
x_{18}	controls	length of brow	.50	.30	1.00
x_{19}	controls r	radius of ear	.50	.10	1.00
x_{20}	controls	nose width	.10	.10	.20

Few of the facial descriptions are entirely accurate. Most of the facial features are controlled by the data associated with the feature and the data associated with other features. For example, the true face width is a function not only of h^* but also of θ^*; mouth length depends on a_m and also on w_m.

The ranges of the facial features have been adjusted so that the faces look more "human" and so that all the features are observable. The eye size has been set so that the pupils can be seen; the mouth length set so that curvature is visible. It is important that all features be observable; that the faces possess human-like features is a matter of preference and appropriateness. It may be that the use of human-like features will contribute to the interpretation of one set of data but not to another.

II. USING THE PROGRAM

To create a Chernoff face an assignment of the data dimensions to the facial features is made. This assignment may be made at random or deliberately. Some users prefer the random assignment to reduce subjective elements, others deliberately employ perception of facial characteristics in the assignment. Thus, a measure of success or failure may be associated with mouth curvature and a measure of liberal/conservative stance with pupil position (looking to the left or right).

Once the assignment is made, low and high data values of each data dimension are determined. The actual value of the data variable will be linearly mapped from the data range into the facial feature range. (It is sometimes advantageous to transform the data by logs or powers before carrying out this step.) When all data ranges are set, the data can be mapped into facial feature ranges and a face which represents the input data drawn.

The data ranges should be set carefully. If the range is set too small, it will not include all of the actual data values; if it is set too large the data values may be too close relative to the range and a loss of distrimination in the facial feature may result. Figure 3 shows the results of data values outside of the mouth length and nose length ranges. Transformed data may spread the actual values more or less uniformly throughout the range and this can lead to increased differentiation. Only knowledge of the data can help one decide what to do.

A listing of the control cards, the computer program, DRFACE, and the data from a study on oil companies involved in outer continental shelf leasing and drilling is contained in the Appendix. Note that the READ statements for the facial feature - data variable assignment allow easy reassignment. This is particularly helpful when the data are to be viewed from several perspectives.

The user specifies one input and two output format statements. The two output formats allow neater display. The actual output for the oil company example follows the data in the Appendix. The facial features chosen, their ranges, the corresponding data variables and their user-given ranges are printed first. Then a three-line set of information for each face drawn is printed. The first line gives the scaled facial data and the third line gives the number of the corresponding facial feature.

From the output data we see that the facial feature "mouth curvature" is associated with the data variable "Royalty per Production Year." (See Table IV for a definition of the latter term.) The low and high values of "Royalty per Production Year" are assigned on the input cards; the low and high values of "mouth curvature" are assigned in DRFACE in the DATA FEAT statement. Figure 4 contains the faces drawn for this example.

At LASL DRFACE is run on both the CDC 7600 and CDC 6600 computers. The former version is the program given in the Appendix. The faces are produced on a Calcomp model 565 plotter. The 6600 version is used in conjunction with a Tektronix 4000 Series CRT screen terminal and a film recorder.

III. OTHER TECHNIQUES FOR DISPLAYING MULTIVARIATE DATA

There are many ways to display multivariate data. Table II presents the data for the oil company example. (The variables are described in Table IV). A look at this table conveys very little information. Figure 4 in the previous section presents the data by means of Chernoff faces.

The same data will now be displayed using Andrews' sine curves and figures called metroglyphs. D. F. Andrews (2) has suggested mapping multidimensional data into trignometric functions on $[-\pi,\pi]$ in the following way

$$(x_1, x_2, \ldots, x_k) \rightarrow \frac{1}{\sqrt{2}} x_1 + x_2 \sin(t) + x_3 \cos(t) + x_4 \sin(2t) + x_5 \cos(2t) + \ldots ,$$

This function is then plotted so that each multidimensional point produces a curve. The curves are viewed and those that lie close together represent clusters. The results for the oil company data are presented in Fig. 5. I find this figure hard to interpret. Shell appears to be different, but the rest are too intertwined. Sometimes plotting principal components rather than the data improves the picture. It did not help in this case.

The first seven variables of the data in Table II are plotted in Fig. 6. These shapes are called metroglyphs (3). These are typical of many other types of multidimensional data display techniques which use circles, rays and location within an area to display the data. This

TABLE II. Data on 15 Variables for 10 Oil Company Groups

Company Group Name	Variable														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ARCO	.56	1.1	.78	306	49	10	4.5	.38	66	62	.11	174	.84	2.8	35
UNION	.53	1.2	.49	203	47	4	4.2	1.22	103	99	.19	527	.98	8.5	38
GETTY	.54	1.0	.32	197	31	11	4.0	.67	51	57	.11	160	.38	2.8	26
MOBIL	1.21	2.8	.50	211	50	8	3.9	1.04	68	78	.06	339	.81	4.0	25
TEXACO	1.16	2.7	.56	176	66	8	7.8	.31	56	50	.04	277	.91	2.5	34
CHEVRON	.84	1.2	1.16	378	70	13	5.8	.70	197	141	.17	355	.50	1.6	32
GULF	1.01	2.2	.67	219	65	11	4.1	1.53	338	235	.23	481	.83	2.9	37
AMOCO	.66	1.3	.66	258	53	8	7.3	.45	37	44	.07	213	.31	2.7	30
SHELL	.97	1.7	1.59	336	95	13	3.6	1.90	430	378	.39	656	.38	2.7	54
EXXON	1.44	2.9	1.02	250	84	8	5.2	.99	276	199	.14	609	.58	4.3	36

particular code for metroglyphs was written by Herbert T. Davis, Jr., of Sandia Laboratories, Albuquerque, New Mexico.

The Andrews' sine curves, the metroglyphs and the Chernoff face are three different ways of graphically representing multidimensional data. Some of the advantages and disadvantages of each will be discussed in the next section.

IV. ADVANTAGES AND DISADVANTAGES OF USING CHERNOFF FACES

Each of the techniques used to display multidimensional data has advantages and disadvantages associated with its use. The Chernoff face method has several distinct advantages over other representatinal techniques such as those presented in the previous section.

First, faces are easily recognized and described. We grow up studying faces and learning to recognize different facial expressions. Professor Chernoff has indicated to me that he chose faces over, say houses, because of our ability to differentiate among the former. Differentiation among metroglyphs or Andrews' sine curves is more difficult. It is not even clear how to describe similarity of sine curves.

When a face is presented we can rely on a commonality of language in our discussions. We speak of nose length or ear height and there is no confusion. Metroglyphs can be described, but not quite as easily. Both of these have the advantage of linking individual data variables with figure characteristics. However, this linkage may not always be meaningful.

This leads to a second advantage of using Chernoff faces. We are able to link facial characteristics with the physical meaning of the variables. The smile can be used to represent a "success/failure" variable, the eyes can represent a "slyness" variable or a political stance, the forehead may represent intelligence as was done by Lt. Gerald Lake in a study here at the Naval Postgraduate School. Research on the perception of facial features is shedding light on appropriate uses.

Unfortunately this may make the use and interpretation of faces more subjective. But is subjectivity entirely bad? I do not think so and list subjectivity as the third advantage of using Chernoff faces. The subjectivity is obvious and this distinguishes the face methodology from other techniques. If we are using the faces for clustering, the clusters formed will be influenced by the facial feature-data variable assignment and by the biases of the viewer. If we use a computer package, the choice of clustering algorithm is a subjective choice. Unfortunately, in the latter case it is all too easy to think of the results are arrived at objectively. This is not likely to happen with the face usage. The metroglyph type representations appear objective, but I'm not sure yet. Do we know that we will get the same clustering no matter how the figures are rotated? The Andrews' sine curves will vary with different orderings of the input data.

It must be remembered that there is no universally accepted correct and true method to arrive at clusters. The faces are not being proposed as a method of arriving at final decisions, but rather as a means of studying the data. If the subjectivity of the faces causes the user to be more careful in his or her conclusions, that is fine. If the apparent objectivity of a technique caused the user to treat the technique as final, that would be a disadvantage.

The fourth and final advantage I will give for using faces is that it is possible to concentrate on subsets of the data variables without redoing the graphics. We might want to concentrate on the variables associated with the eyes and ears and then concentrate on the variables associated with the ears and mouth. This concentration is virtually impossible if one uses Andrews' sine curves, and not as convenient when using metroglyphs. The latter suffers from a description problem.

In spite of the above pulses for using Chernoff faces, there are some minuses. Perhaps the first disadvantage to using the faces is that a plotting device is required if one is to draw a standard Chernoff face. I say standard because of a paper (4) by Turner and Tidmore presented at the 1977 Chicago meeting of the American Statistical Association. They demonstrated how Chernoff-type faces can be drawn with a line printer.

A second disadvantage is that a new chapter on the use of Chernoff faces to deceive could be added to Darrell Huff's <u>How to Lie With Statistics.</u> (5) The faces can be abused. However, if we refused to use any technique which can be misused, there would be little left.

A more serious problem with the Chernoff faces is that the built-in dependencies among facial features may distort the data representation enough to cause erroneous impressions. In Section 5, I will discuss this topic more completely. It should be noted that even if all the facial features are independent, there is no guarantee, in fact it is rather unlikely, that the total face will be viewed as a union of 20 different, independent variables.

The final point I'd like to make is that as the number of entities to be represented increases, severe difficulties may occur in actually viewing the faces. This will be particularly true if the faces are similar. If there are two or three very different classes, there won't be much difficulty, if any. This same problem will occur with metroglyphs, Andrews' sine curves and most other representational modes when the number of entities is large.

A similar problem occurs if we try to use all 20 dimensions in the Chernoff faces. Viewing becomes difficult. Fifteen variables is a good maximum. Metroglyph type representations also suffer if the number of dimensions gets too large.

V. FACIAL FEATURE DEPENDENCIES

There is a potentially serious problem involved in using the faces to represent data. While some of the facial features depend only on the input data for the corresponding data variable, other facial features are interrelated to some extent. Face height, the three facial eccentricities, eye slant, ear level and ear size are in the former class; most of the remaining features are in the latter. Pupil position does depend on other facial features only to guarantee that the pupil remains in the eye. The mouth structure, however, depends on face height and width, ear level, lower face eccentricity and nose length, as well as on the three mouth parameters. Eye height depends on nose length and face height; eye separation on the upper face eccentricity and face height. These dependencies occur in order to insure proper positioning of the facial features.

The results of the dependencies can be deceiving. Figure 7 shows eight faces in which all parameters except ear level, nose length and lower eccentricity remain constant. Table III identifies the cases. Notice the effect of these three facial features on the mouth length and forehead.

In Chernoff's original program, the faces were normalized so that both face height and width were constant. The normalization reduces the dependencies, but does not eliminate them. It does, however, essentially remove the face height and width variables from consideration. The program DRFACE does not contain the normalization.

Restriction of the ranges of the facial features reduces the dependencies somewhat. Not using face height, the eccentricities of the upper and lower ellipses and nose length would help greatly but would also cost in terms of lost variables. Loss of nose length is particularly undesirable. Perhaps one solution to the problem is to identify non-overlapping regions for the features and then restructure

TABLE III. Facial Dependency Parameters

	Facial Feature		
Case	Nose Length	Lower Eccentricity	Ear Level
1	.15	.50	.35
2	.15	.50	.65
3	.15	1.00	.35
4	.15	1.00	.65
5	.40	.50	.35
6	.40	.50	.65
7	.40	1.00	.35
8	.40	1.00	.65

DRFACE so that the facial features must lie within these regions. See Fig. 8. Mathematical dependencies would thus be removed. (Perception dependencies may still exist.) There may also be some merit to setting the upper and lower face eccentricities to 1. This will make the face circular.

VI. APPLICATIONS TO CHERNOFF FACES AT LASL

The main application I have made of the faces technique in the past has been to represent data on some of the major oil and gas companies involved in offshore leasing. The ten oil company groups are described in Table IV. (The Arbitrary Company Code (ACC) is a designation given the companies by the Conservation Division, U. S. Geological Survey, Denver, CO.) The variables considered are contained in Table V.

Myrle Johnson of LASL has used the faces program to describe energy related variables on a state by state basis (6). Her paper contains many other examples of the use of graphics to represent data.

Presently I am collaborating with James McFarland and Laird Landon of the University of Houston to use faces to represent quarterly data on nine banks in the Houston area. Faces have been drawn with both random and planned assignment of facial features to data variables. The faces

TABLE IV. Arbitrary Company Codes (ACC)

ACC	NAME	COMPANIES
2	ARCO	AtlanticRichfield, Richfield Oil, Sinclair, B. B. Barber, Barber Oil Exploration, Royal Gorge Company
3	UNION	Union, Pure Oil, Pure Transportation Company
5	GETTY	Getty, Skelly
39	MOBIL	Mobil, Magnolia Petroleum
40	TEXACO	Texaco, Texaco Seaboard
78	CHEVRON	Chevron, California Company, Standard Oil of Texas
112	GULF	Gulf, British Americal Oil
114	AMOCO	Amoco, Midwest Oil, Standolind, Pan American
117	SHELL	Shell, Shell P/L Corporation
276	EXXON	Exxon, Humble, Exxon Pipeline Company

are to be presented to classes at the University for clustering by students.

Plans are being made to use faces to represent the changing chemical content of water from 17 wells in Los Alamos County. Also it is possible that faces may be used to display the results of an employee attitude survey conducted at LASL the week of February 6-10, 1978. It is proposed to draw one or two faces for each of the Laboratory's 18 divisions.

I would like to discuss briefly one new possible application of Chernoff faces. All the applications discussed so far use faces to display attribute data of some population of interest. I am attempting to see if it is possible to use faces in distributional studies.

The problem being considered is this. Suppose we have a random sample of size 15 from either a normal distribution, N(0,1), or a rectangular distribution on $(-\sqrt{3},\sqrt{3})$. Both of these distributions have

TABLE V. Description of External Variables

1	Net Bonus	Total net dollars paid (in billions)
2	Excess $/Lease	Average gross dollars paid above 2nd highest bid (in millions).
3	Net acreage	Total net acreas leased (in millions)
4	No. leases won	Number of leases won
5	Avg. Ownership	Average percent of ownership of leases
6	Pct. Prod. Leases won	Percentage of leases, ultimately found to be producing, won by the company
7	Avg. Yrs. to Prod.	Average number of years between sale and first production (production lag)
8	Net Gas Prod.	Net gas production (in trillion cubic feet)
9	Net Liq. Prod.	Net liquid production (in millions of barrels)
10	Net Royalty	Net royalty paid to government (in millions)
11	Royalty/Bonus	Net royalty/number of years of production (in thousands)
13	R**2:Roy/Pyr.	Square of correlation coefficient from multiple linear regression of royalty/prod. yr. on bidding data, production lag and years of production (for producing leases only)
14	Roy/Pyr/Pr.Ac.	Royalty/production year/producing acre (in dollars)
15	Pct. Leases Term-ted	Percentage of owned leases terminated

zero mean and unit variance.) We would like to determine the distribution from which the sample came.

Two approaches are being investigated. In the first, the sample is ordered and the order statistics obtained. Each of the 15 order statistics is assigned to a facial feature, and a face is drawn. This face is then compared to the nominal faces drawn using the same facial feature assignment and the expected values of the order statistics for the normal and rectangular distributions. The distribution whose nominal face is most similar to the sample face is chosen as the parent distribution.

The second approach is to use the face to summarize sample and test statistics and then to compare the sample face to the population faces. Some statistics to be considered are skewness and kurtosis other sample moments, values of the χ^2 and Kolmogorov-Smirnov statistics and perhaps some order statistics or functions of such.

VII. CONCLUSIONS

The Chernoff face technique is one of several available graphical techniques used to display and analyze multidimensional data. When properly employed, it provides useful insight into the nature of the data and has some important advantages over the other graphical techniques. The main criticism of the technique, its subjectivity, can actually be considered as a positive feature rather than as a drawback. The problem of facial feature dependencies can be overcome. In recent years there have been many interesting applications of Chernoff faces in conjunction with cluster analysis, outlier detection methods, distributional studies, and time-series analysis. The Chernoff faces technique will continue to be an effective tool in exploratory multivariate data analysis.

REFERENCES

1. Chernoff, H. (1971). "The Use of Faces to Represent Points in n-Dimensional Space Graphically," Technical Report No. 71, Department of Statistics, Stanford University, Stanford, CA.
2. Andrews, D. F. (1972). "Plots of High Dimensional Data," Biometrics, 28, 125-136.
3. Gnanadesikan, R. (1977). "Methods for Statistical Data Analysis of Multivariate Observations," John Wiley & Sons, NY
4. Turner, D. W., Tidmore, F. E. (1977). "Clustering with Chernoff-Type Faces," Proceedings of the Statistics Section, American Statistical Association Annual Meeting, Chicago, IL.
5. Huff, D. (1954) "How to Lie with Statistics," W. W. Norton & Co.,
6. Lohrding, R. K., Johnson, M. M., Whiteman, D. E. (1978). "Computer Graphics For Extracting Information from Data," to be presented at the 11th Annual Symposium on the Interface. Computer Science and Statistics, March 6-7, Raleigh, N.C.

APPENDIX

```
$FUN(C=S)
$OPEN(FS=IMAGE,SCT=10000)
$OPEN(FS=FILM ,SCT=10000)
$LDGO.
$AFSREL(FS=FSET12,ADISP=TAPE6,POSDEN=556,POSMT=LA350L00)
$FM.
      PROGRAM  DRFACE(INP,OUT,      FSET5=INP,FSET6=OUT,FSET12)
      DIMENSION XFACE(1000),YFACE(1000),YSAME(201),XNOSE(51),
     1YNOSE(51),XMOUTH(51),YMOUTH(51),XLEYE(80),XREYE(80),YEYES(8
     2XLBROW(41),PUPILX(2),PUPILY(2),XRBROW(41),YBROWS(41),
     3Y(40),RANGEY(39),AI(39),BI(39),BIMAI(39),
     4DATA(39),AMODES(200),RAN(2,2), TEMP(39),
     5FEAT(20,2)
      REAL LB, LHSRHS(400),MINY(39),MAXY(39),ID1(1000),L,LSQ
      LOGICAL  IFMT(18),OFMT1(18),OFMT2(18)
      DIMENSION IFOX(39),IRAND(39),IVAR(2)
       DATA  PI/3.141593/,BLANK/1H /
      DATA IFOX/39*0/,NVAR/20/
      DATA Y/40*0.0/,IRAND/39*1/
C     DATA FEAT/.2,.35,.5,.5,.5,.15,.2,-4.,.3,0.,
      DATA FEAT/.57,.35,.5,.5,.5,.15,.2,-4.,.3,0.,
     1.3,.2,.4,.2,.2,.6,0.,.3,.1,.1,
C    2.7,.65,1.,1.,1.,.4,.8,4.,1.,.3,.8,.6,.8,1.,.8,1.,1.,1.,1.,.
     2 68,.65,1.,1.,1.,.4,.8,4.,1.,.3,.8,.6,.8,1.,.8,1.,1.,1.,1.,
      D TA D1,D2/1.,1./
      DATA DATA/0.6, 0.5, 0.5, 0.5, 1.0,0.25, 0.5, 0.0, 0.5, 0.1,
     1          0.7,0.5,0.6,0.5,0.5,0.8,0.5,0.5,0.5,0.1/
      D TA CAPH/1.0/,XINIT/0.0/,YINIT/0.0/
      INTEGER    NU,NL,NFACE,NNOSE,NMOUTH,NEYES,NBROWS
     *,ISTR,IDENT(100)
      DATA NU,NL,NNOSE,NMOUTH,NEYES,NBROWS/400,400,51,51,80,41/
      READ 1,NPLOTS,NFIXED
    1 FORMAT(19I4)
C
C
      WRITE(6,400)
C  READ AND PRINT CONTROL CARDS.
 400    FORMAT( 35X,*CONTROL CARDS READ*//
     $*      FACIAL FEATURE*,25X,*      EXTERNAL VARIABLE*//
     L* NO.  NAME                       LOW     HIGH       NO.  NAME
     $     LOW     HIGH*/ )
      DO 16 IFEAT=1,NFIXED
      READ 410,IVAR(1),TP1,TP2,IVAR(2),TP3,TP4,(RAN(2,J),J=1,2)
  4 0 FORMAT(I4,2A10/40X,I4,2A10,2F8.2)
      JJ=IVAR(1)
      RAN(1,1)=FEAT(JJ,1)
      RAN(1,2)=FEAT(JJ,2)
      WRITE(6,411)(IVAR(1),TP1,TP2,(RAN(1,J),J=1,2),IVAR(2),TP3,T
     1,RAN(2,J),J=1,2))
  411 FORMAT(I4,2A10,2F8.2,5X,I4,2A10,2F8.2)
      KK=IVAR(2)
      IFOX(KK)=IVAR(1)
      KP=IVAR(1)
      BI(KP)=(RAN(1,2)-RAN(1,1))/(RAN(2,2) - RAN(2,1))
   1  AI(KP)=RAN(1,1) - RAN(2,1)*BI(KP)
C
      READ(5,405)IFMT
C     READ IN FORMAT FOR DATA     AND OUT FORMAT FOR DATA
      READ(5,405)OFMT1,OFMT2
  405 FORMAT(18A4)
C
      IPLOT=0
      CALL PLOTS(12)
      CALL PLTZ(0.,-12.,-3)
      YYY=YINIT +1.25
       XXX=XINIT
      CALL PLTZ(0.,1.25,-3)
      NYP=4
COMPUTE VERTICAL    ( Y DIRECTION) NUMBER OF PLOTS
      DO 50 JPLOT4=1,NPLOTS,NYP
 446    FORMAT(10F7.1)
      XXX=XXX+3.
      YYY=YYY+7.50
      CALL PLTZ(3.,7.5,-3)
      ITOP=1
      JEND =MINO(NPLOTS,JPLOT4+NYP-1)
```

```
      DO 49 JPLOT1=JPLOT4,JEND
      READ(5,IFMT)XID1,(Y(I),I=1,NFIXED)
      WRITE(6,1)
      WRITE(6,1)
      WRITE(6,OFMT1)XID1,(Y(I),I=1,NFIXED)
    2 FORMAT(9F8.2)
      DO 15 IVB=1,NFIXED
      IF(IFOX(IVB).EQ.0)GO TO 15
      KKK=IFOX(IVB)
      DATA(KKK)=AI(KKK)+BI(KKK)*Y(IVB)
   15 CONTINUE
      IF(ITOP.EQ.1)GO TO 17
      YYY=YYY-2.50
      CALL PLTZ(0.,-2.5,-3)
   17 ITOP=0
      DO 13 J=1,NFIXED
      JTP=IFOX(J)
   13 TEMP(J)=DATA(JTP)
      WRITE(6,OFMT1)BLANK,(TEMP(J),J=1,NFIXED)
      WRITE(6,OFMT2)BLANK,(IFOX(J),J=1,NFIXED)
C
C
      HSTAR=.5*(1.0+DATA(1))*CAPH
      THSTAR=(2.0*DATA(2)-1.0)*PI*0.25
      SMALLH=.5*(1.0+DATA(3))*CAPH
      X0=HSTAR*COS(THSTAR)
      Y0=HSTAR*SIN(THSTAR)
C
C     COMPUTE  FACE
      CU=.5*(SMALLH+Y0-X0**2/(DATA(4)**2*(SMALLH-Y0)))
      BU=SMALLH-CU
      AU=DATA(4)*BU
      BUSQ=BU**2
      CL=.5*(-SMALLH+Y0-X0**2/(DATA(5)**2*(-SMALLH-Y0)))
      BL=SMALLH+CL
      AL=DATA(5)*BL
      BLSQ=BL**2
      XMAX=X0
      NFACE=NU+NL
C
      NUP1=NU+1
      YSAME(1)=Y0
      LHSRHS(1)=-X0
      NSTEP=NU/2
      NSTPP1=NSTEP+1
      YSAME(NSTPP1)=SMALLH
      LHSRHS(NSTPP1)=0.0
      STPSIZ=(SMALLH-Y0)/NSTEP
      ISTOP=NSTEP-1
C
      DO 5 I=1,ISTOP
      IP1=I+1
      YSAME(IP1)=Y0+I*STPSIZ
      NUMI=NUP1-I
      XPLUS=DATA(4)*SQRT(BUSQ-(YSAME(IP1)-CU)**2)
      IF(XPLUS.GT.XMAX)XMAX=XPLUS
      LHSRHS(IP1)=-XPLUS
      LHSRHS(NUMI)=XPLUS
    5 CONTINUE
C
      XFACE(1)=LHSRHS(1)
      YFACE(1)=YSAME(1)
      NUP2=NU+2
      DO 6 I=2,NSTEP
      XFACE(I)=LHSRHS(I)
      YFACE(I)=YSAME(I)
      IX2=NUP2-I
      XFACE(IX2)=LHSRHS(IX2)
    6 YFACE(IX2)=YSAME(I)
      XFACE(NSTPP1)=LHSRHS(NSTPP1)
      YFACE(NSTPP1)=YSAME(NSTPP1)
C
```

```
      YSAME(1)=YO
      LHSRHS(1)=XO
      NLP1=NL+1
      YSAME(NSTPP1)=-SMALLH
      LHSRHS(NSTPP1)=0.0
      STPSIZ=(YO+SMALLH)/NSTEP
      DO 7 I=1,ISTOP
      IP1=I+1
      NLMI=NLP1-I
      YSAME(IP1)=YO-I*STPSIZ
      XPLUS=DATA(5)*SQRT(BLSQ-(YSAME(I)-CL)**2)
      IF(XPLUS.GT.XMAX)XMAX=XPLUS
      LHSRHS(IP1)= XPLUS
      LHSRHS(NLMI)=-XPLUS
    7 CONTINUE
C
      NLP2=NL+2
      XFACE(NUP1)=LHSRHS(1)
      YFACE(NUP1)=YSAME(1)
      DO 8 I=2,NSTEP
      XFACE(NU+I)=LHSRHS(I)
      YFACE(NU+I)=YSAME(I)
      IX2=NLP2-I
      XFACE(NU+IX2)=LHSRHS(IX2)
      YFACE(NU+IX2)=YSAME(I)
    8 CONTINUE
      XFACE(NU+NSTPP1)=LHSRHS(NSTPP1)
      YFACE(NU+NSTPP1)=YSAME(NSTPP1)
      XMIN=-XMAX
      YMAX=SMALLH
      YMIN=-SMALLH
C
C     COMPUTE  NOSE
      AN=SMALLH*DATA(6)
      XNOSE(1)=0.0
      XNOSE(2)=SMALLH*DATA(20)
      XNOSE(3)=-XNOSE(2)
      YNOSE(1)=AN
      YNOSE(2)=-AN
C
    9       CONTINUE
C     COMPUTE  MOUTH
      YM=-SMALLH*(DATA(6)+(1.0-DATA(6))*DATA(7))
      XOFYM=DATA(5)*SQRT(BLSQ-(YM-CL)**2)
      AX8=SMALLH/ABS(DATA(8))
      AM=DATA(9)*AMIN1(XOFYM,AX8)
      NSTEP=NMOUTH/2
      NMP1=NMOUTH+1
      YMOUTH(NSTEP+1)=YM
      XMOUTH(NSTEP+1)=0.0
      STPSIZ=AM/NSTEP
      X8SQ=(SMALLH/DATA(8))**2
      HBY8=AX8
      IF(DATA(8).LT.0.0)SIGN=-1.0
      IF(DATA(8).GT.0.0)SIGN=1.0
      DO 11 I=1,NSTEP
      XPLUS=-AM+(I-1)*STPSIZ
      XMOUTH(I)=XPLUS
      NMMI=NMP1-I
      XMOUTH(NMMI)=-XPLUS
      YMOUTH(I)=YM+SIGN*(HBY8  -SQRT(X8SQ-XPLUS**2))
   11 YMOUTH(NMMI)=YMOUTH(I)
C
C     COMPUTE  EYES
      YE=SMALLH*(DATA(6)+(1.0-DATA(6))*DATA(10))
       XOFYE=DATA(4)*SQRT(ABS(BUSQ-(YE-CU)**2))
      XE=XOFYE*(1.0+2.0*DATA(11))*0.25
      THETA=(2.0*DATA(12)-1.0)*PI*0.2
      X13=DATA(13)
      L=DATA(14)*AMIN1(XE,XOFYE-XE)
      LSQ=L**2
      SINTH=SIN(THETA)
      COSTH=COS(THETA)
      R=L/SQRT(COSTH**2+SINTH**2/X13**2)
      PUPILX(1)=-XE+R*(2.0*DATA(15)-1.0)
      PUPILX(2)=XE+R*(2.0*DATA(15)-1.0)
      PUPILY(1)=YE+R*(2.0*DATA(15)-1.0)*TAN(THETA)
      RPUP=DATA(13)*DATA(14)/10.
```

```
C
      NSTEP=NEYES/4
      STPSIZ=L/NSTEP
      I1=1
      I2=NSTEP+1
      I3=2*NSTEP+1
      I4=3*NSTEP+1
      U=0.0
      V=X13*L
      XSTAR=-V*SINTH
      YSTAR= V*COSTH
      XX=XE+XSTAR
      YY=YE+YSTAR
      XREYE(I2)=XX
      YEYES(I2)=YY
      XLEYE(I2)=-XX
      XX=XE-XSTAR
      YY=YE-YSTAR
      XREYE(I4)=XX
      XLEYE(I4)=-XX
      YEYES(I4)=YY
      U=L
      XSTAR=U*COSTH
      YSTAR=U*SINTH
      XX=XE+XSTAR
      YY=YE+YSTAR
      XREYE(I3)=XX
      XLEYE(I3)=-XX
      YEYES(I3)=YY
      XX=XE-XSTAR
      YY=YE-YSTAR
      XREYE(I1)=XX
      XLEYE(I1)=-XX
      YEYES(I1)=YY
      I1=I2
      I3=I4
      ISTOP=NSTEP-1
   10    CONTINUE
      DO 12 I=1,ISTOP
      U=I*STPSIZ
      V=X13*SQRT(LSQ-U**2)
      XSTAR=U*COSTH-V*SINTH
      YSTAR=U*SINTH+V*COSTH
      XX=XE+XSTAR
      YY=YE+YSTAR
      I2=I2+1
      I4=I4+1
      XREYE(I2)=XX
      XLEYE(I2)=-XX
      YEYES(I2)=YY
      XX=XE-XSTAR
      YY=YE-YSTAR
      XREYE(I4)=XX
      XLEYE(I4)=-XX
      YEYES(I4)=YY
      XSTAR=U*COSTH+V*SINTH
      YSTAR=U*SINTH-V*COSTH
      I1=I1-1
      I3=I3-1
      XX=XE-XSTAR
      YY=YE-YSTAR
      XREYE(I1)=XX
      XLEYE(I1)=-XX
      YEYES(I1)=YY
      XX=XE+XSTAR
      YY=YE+YSTAR
      XREYE(I3)=XX
      XLEYE(I3)=-XX
      YEYES(I3)=YY
   12 CONTINUE
C
```

```
C      DRAW EYEBROWS
       YB=YE+2.0*(0.3 + DATA(16))   *L*X13
       THSTST=THETA+PI*(2.0*DATA(17)-1.0)*0.2
       COSTH=COS(THSTST)
       SINTH=SIN(THSTST)
       LB=R*(2.0*DATA(18)+1.0)*0.5
       XX=LB*COSTH+XE
       YY=LB*SINTH+YB
       XRBROW(1)=XX
       XLBROW(1) =-XX
       YBROWS(1) =YY
       XX=-LB*COSTH+XE
       YY=-LB*SINTH+YB
       XRBROW(2) =XX
       XLBROW(2) =-XX
       YBROWS(2) =YY
C
COMPUTE EARS
       REAR=(1.0+DATA(19))*SMALLH*.1
       CEAR=HSTAR+REAR
       EARX=CEAR*COS(THSTAR)
       EARY=CEAR*SIN(THSTAR)
C
C      SET PARAMETERS
C
  61      CONTINUE
       CALL SYMBOL(-1.25,-1.25,.2,XID1,0.,10)
C
  62      CONTINUE
C      DRAW FACE
       XV=0.
       YV=0.
       CALL LINE(XFACE,XV,D1,YFACE,YV,D2,NU,1,0,55B,0)
       CALL LINE(XFACE(NUP1),XV,D1,YFACE(NUP1),YV,D2,NL,1,0,55B,0)
C
  63      CONTINUE
C      NOSE
       CALL PLTZ(XNOSE(1),YNOSE(1),3)
       CALL PLTZ(XNOSE(2),YNOSE(2),2)
        CALL PLTZ(XNOSE(3),YNOSE(2),2)
        CALL PLTZ(XNOSE(1),YNOSE(1),2)
C
  64      CONTINUE
C      MOUTH
       CALL LINE(XMOUTH,XV,D1,YMOUTH,YV,D2,NMOUTH,1,0,55B,0)
  65      CONTINUE
C      EYES
       CALL LINE(XLEYE,XV,D1,YEYES,YV,D2,NEYES,1,0,1H )
       CALL LINE(XREYE,XV,D1,YEYES,YV,D2,NEYES,1,0,1H )
C
  66      CONTINUE
C      EYEBROWS
       CALL PLTZ(XRBROW(1),YBROWS(1),3)
       CALL PLTZ(XRBROW(2),YBROWS(2),2)
       CALL PLTZ(XLBROW(1),YBROWS(1),3)
       CALL PLTZ(XLBROW(2),YBROWS(2),2)
C
  67      CONTINUE
C      PUPILS
       CALL CIRCLE(PUPILX(1),PUPILY(1),RPUP,20)
       CALL CIRCLE( PUPILX(2),PUPILY(1),RPUP,20)
C      DRAW EARS
       CALL CIRCLE(EARX,EARY,REAR,20)
       CALL CIRCLE(-EARX,EARY,REAR,20)
   49  CONTINUE
   50  CONTINUE
       WRITE(6,4051) IFMT
 4051  FORMAT(*0INPUT FORMAT IS *,18A4)
       WRITE(6,4052) OFMT1
 4052  FORMAT(*0OUTPUT FORMAT IS *,18A4)
       CALL PLTZ(0.,0.,999)
C
C
C
       STOP
       END
```

```
      SUBROUTINE CIRCLE(XO,YO,RAD,NPTS)
      DELTH=6.283185/FLOAT(NPTS)
      THETA=0.0
      XX=XO+RAD
      YY=YO
      CALL PLTZ(XX,YY,3)
      DO 5 I=1,NPTS
      THETA=THETA+DELTH
      XX=XO+RAD*COS(THETA)
      YY=YO+RAD*SIN(THETA)
      CALL PLTZ(XX,YY,2)
    5 CONTINUE
      RETURN
      END
$FM.
```

```
10 15
 1 FACE WIDTH
18 BROW LENGTH
 3 FACE HEIGHT
11 EYE SEPARATION
15 PUPIL POSITION
 6 NOSE LENGTH
20 NOSE WIDTH
19 EAR DIAMETER
 2 EAR LEVEL
 9 MOUTH LENGTH
12 EYE SLANT
 8 MOUTH CURVATURE
 7 MOUTH LEVEL
10 EYE LEVEL
16 BROW HEIGHT
```

1	NET BONUS (B$)	.5	1.5
2	EXCESS $/LEASE (MM)	1.0	3.0
3	NET ACREAGE (MM)	.3	1.6
4	NO. LEASES WON	175.0	380.0
5	AVG. OWNERSHIP	30.0	100.0
6	PCT.PROD.LEASES WON	0.0	15.0
7	AVG.YRS.TO PROD.	3.0	8.0
8	NET GAS PRO. (TCF)	.3	2.0
9	NET LIQ.PRO. (MMB)	35.0	430.0
10	NET ROYALTY (MM$)	40.0	380.0
11	ROYALTY/BONUS ($)	0.0	.4
12	ROYALTY/PRO.YR (K$)	170.0	660.0
13	R**2: ROY/PYR=F( )	.3	1.0
14	ROY/PYR/PR.AC. ($)	1.5	8.5
15	PCT. [illegible]SES TERM"TED	25.0	60.0

```
(A7,2F4.0,13F5.0)
(1X,A10,16F7.2)
(1X,A10,16I7)
```

ARCO	.56	1.1	.73	306	49	10	4.5	.38	66	62	.11	174	.34	2.8	35
UNION	.53	1.2	.49	203	47	4	4.2	1.22	103	99	.19	527	.98	8.5	38
GETTY	.54	1.0	.32	197	31	11	4.0	.67	51	57	.11	160	.38	2.8	26
MOBIL	1.21	2.8	.50	211	50	8	3.9	1.04	68	78	.06	339	.81	4.0	25
TEXACO	1.16	2.7	.56	176	66	8	7.8	.31	56	50	.04	277	.91	2.5	34
CHEVRON	.84	1.2	1.16	378	70	13	5.8	.70	197	141	.17	355	.50	1.6	32
GULF	1.01	2.2	.67	219	65	11	4.1	1.53	338	235	.23	481	.83	2.9	37
AMOCO	.66	1.3	.66	258	53	8	7.3	.45	37	44	.07	213	.31	2.7	30
SHELL	.97	1.7	1.59	336	95	13	3.6	1.90	430	378	.39	656	.38	2.7	54
EXXON	1.44	2.9	1.02	250	84	8	5.2	.99	276	199	.14	609	.58	4.3	36

FACIAL FEATURE				EXTERNAL VARIABLE			
NO.	NAME	LOW	HIGH	NO.	NAME	LOW	HIGH
1	FACE WIDTH	.20	.70	1	NET BONUS (B$)	.50	1.50
18	BROW LENGTH	.30	1.00	2	EXCESS $/LEASE (MM)	1.00	3.00
3	FACE HEIGHT	.50	1.00	3	NET ACREAGE (MM)	.30	1.60
11	EYE SEPARATION	.30	.80	4	NO. LEASES WON	175.00	380.00
15	PUPIL POSITION	.20	.60	5	AVG. OWNERSHIP	30.00	100.00
6	NOSE LENGTH	.10	.40	6	PCT.PROD.LEASES WON	0.00	15.00
20	NOSE WIDTH	.10	.20	7	AVG.YRS.TO PROD.	3.00	8.00
19	EAR DIAMETER	.10	1.00	8	NET GAS PRO. (TCF)	.30	2.00
2	EAR LEVEL	.30	.60	9	NET LIQ.PRO. (MMB)	35.00	430.00
9	MOUTH LENGTH	.30	1.00	10	NET ROYALTY (MM$)	40.00	380.00
12	EYE SLANT	.20	.60	11	ROYALTY/BONUS ($)	0.00	.40
8	MOUTH CURVATURE	-4.00	4.00	12	ROYALTY/PRO.YR (K$)	170.00	660.00
7	MOUTH LEVEL	.20	.60	13	R**2 ROY/PYR=F( )	.30	1.00
10	EYE LEVEL	0.00	.30	14	ROY/PYR/PR.AC. ($)	1.50	8.50
16	BROW HEIGHT	.60	1.00	15	PCT.LEASES TERM"TED	25.00	60.00

INPUT FORMAT IS (A7,2F4.0,13F5.0)
OUTPUT FORMAT IS (1X,A10,16F7.2)

ARCO	.56	1.10	.73	306.00	49.00	10.00	4.50	.38	66.00	62.00	.11	174.00	.34	2.80	35.00	
	.23	.34	.67	.62	.36	.30	.13	.14	.32	.35	.31	-3.93	.23	.06	.71	
	1	18	3	11	15	6	20	19	2	9	12	8	7	10	16	
UNION	.53	1.20	.49	203.00	47.00	4.00	4.20	1.22	103.00	99.00	.19	527.00	.98	8.50	38.00	
	.21	.37	.57	.37	.35	.18	.12	.59	.35	.42	.39	1.83	.78	.30	.75	
	1	18	3	11	15	6	20	19	2	9	12	8	7	10	16	
GETTY	.54	1.00	.32	197.00	31.00	11.00	4.00	.67	51.00	57.00	.11	160.00	.38	2.80	26.00	
	.22	.30	.51	.35	.21	.32	.12	.30	.31	.33	.31	-4.16	.27	.06	.61	
	1	18	3	11	15	6	20	19	2	9	12	8	7	10	16	
MOBIL	1.21	2.80	.50	211.00	50.00	8.00	3.90	1.04	68.00	78.00	.06	339.00	.81	4.00	25.00	
	.55	.93	.58	.39	.37	.26	.12	.49	.33	.38	.26	-1.24	.64	.11	.60	
	1	18	3	11	15	6	20	19	2	9	12	8	7	10	16	
TEXACO	1.16	2.70	.56	176.00	66.00	8.00	7.80	.31	56.00	50.00	.04	277.00	.91	2.50	34.00	
	.53	.89	.60	.30	.51	.26	.20	.11	.32	.32	.24	-2.25	.72	.04	.70	
	1	18	3	11	15	6	20	19	2	9	12	8	7	10	16	
CHEVRON	.84	1.20	1.16	378.00	70.00	13.00	5.80	.70	197.00	141.00	.17	355.00	.50	1.60	32.00	
	.37	.37	.83	.60	.54	.36	.16	.31	.42	.51	.37	-.98	.37	.00	.68	
	1	18	3	11	15	6	20	19	2	9	12	8	7	10	16	
GULF	1.01	2.20	.67	219.00	65.00	11.00	4.10	1.53	338.00	235.00	.23	481.00	.83	2.90	37.00	
	.45	.72	.64	.41	.50	.32	.12	.75	.53	.70	.43	1.08	.65	.06	.74	
	1	18	3	11	15	6	20	19	2	9	12	8	7	10	16	
AMOCO	.66	1.30	.66	258.00	53.00	8.00	7.30	.45	37.00	44.00	.07	213.00	.31	2.70	30.00	
	.26	.40	.64	.50	.40	.26	.19	.18	.30	.31	.27	-3.30	.21	.05	.56	
	1	18	3	11	15	6	20	19	2	9	12	8	7	10	16	
SHELL	.97	1.70	1.59	336.00	95.00	13.00	3.60	1.90	430.00	378.00	.39	656.00	.38	2.70	54.00	
	.43	.54	1.00	.69	.76	.36	.11	.95	.60	1.00	.59	3.93	.27	.05	.93	
	1	18	3	11	15	6	20	19	2	9	12	8	7	10	16	
EXXON	1.44	2.90	1.02	250.00	84.00	8.00	5.20	.99	276.00	199.00	.14	609.00	.58	4.30	36.00	
	.67	.96	.78	.48	.66	.26	.14	.47	.48	.63	.34	3.17	.44	.12	.73	
	1	18	3	11	15	6	20	19	2	9	12	8	7	10	16	

DRFACE OUTPUT

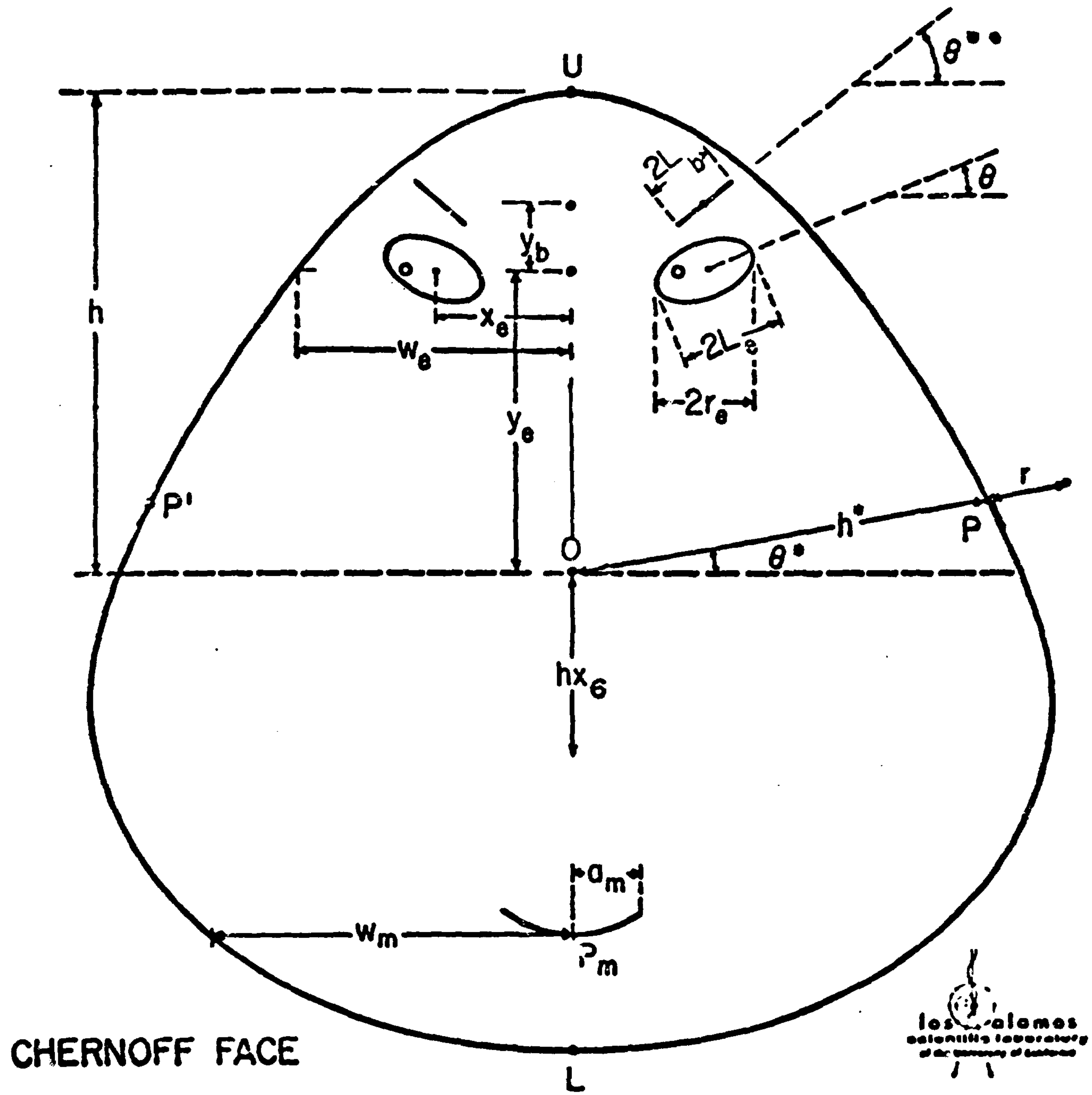


Figure 1. The Original Chernoff Face

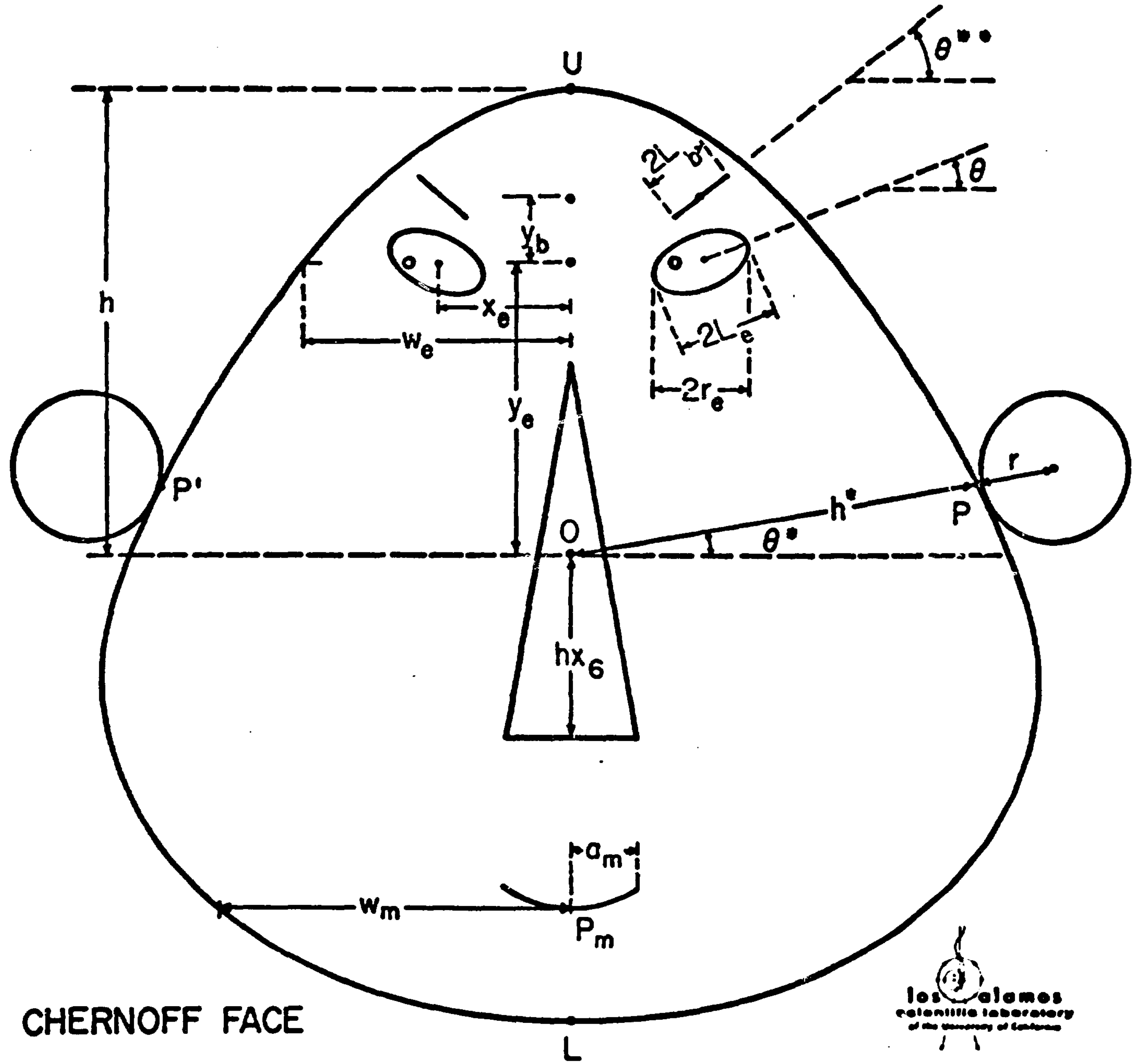
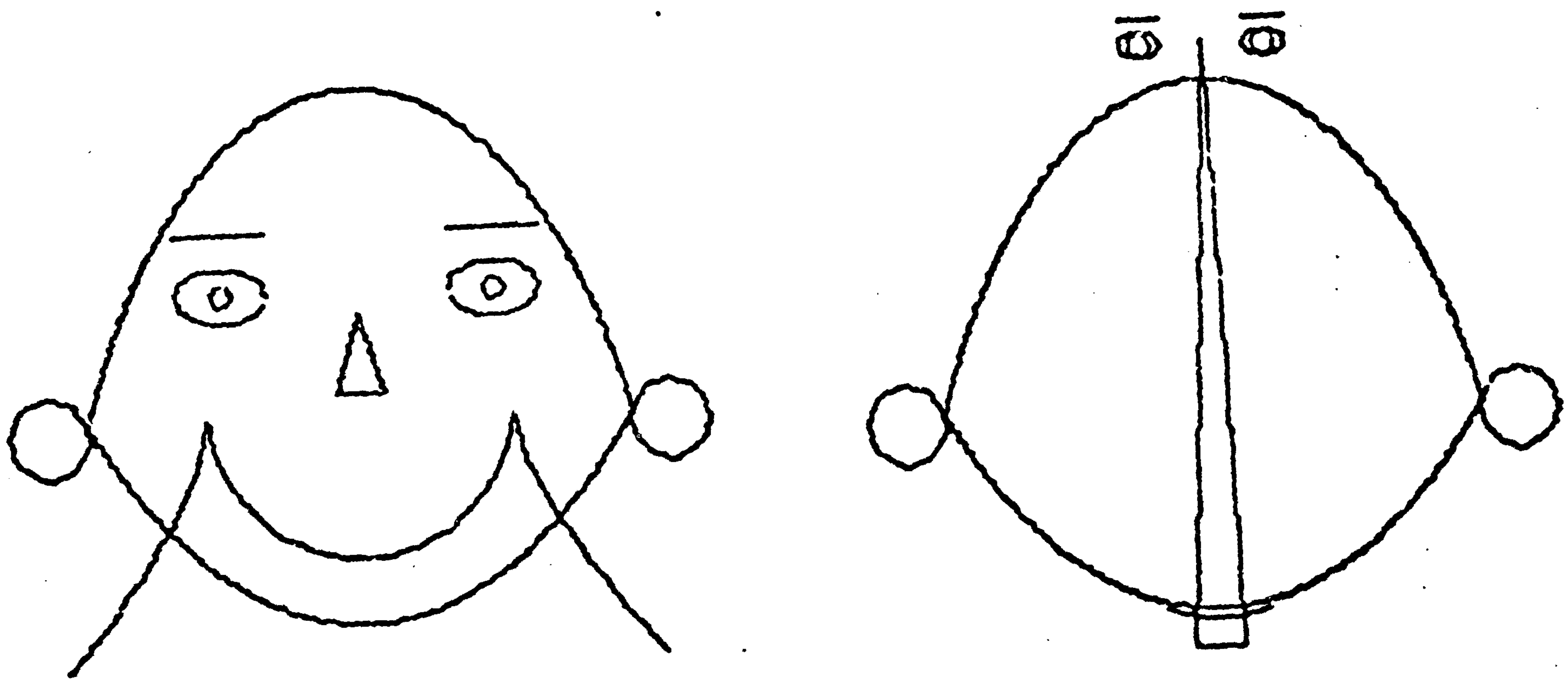


Figure 2. Davis' Chernoff Face



a   b

Figure 3. Effect of mouth length (a) and nose length (b) being out of range

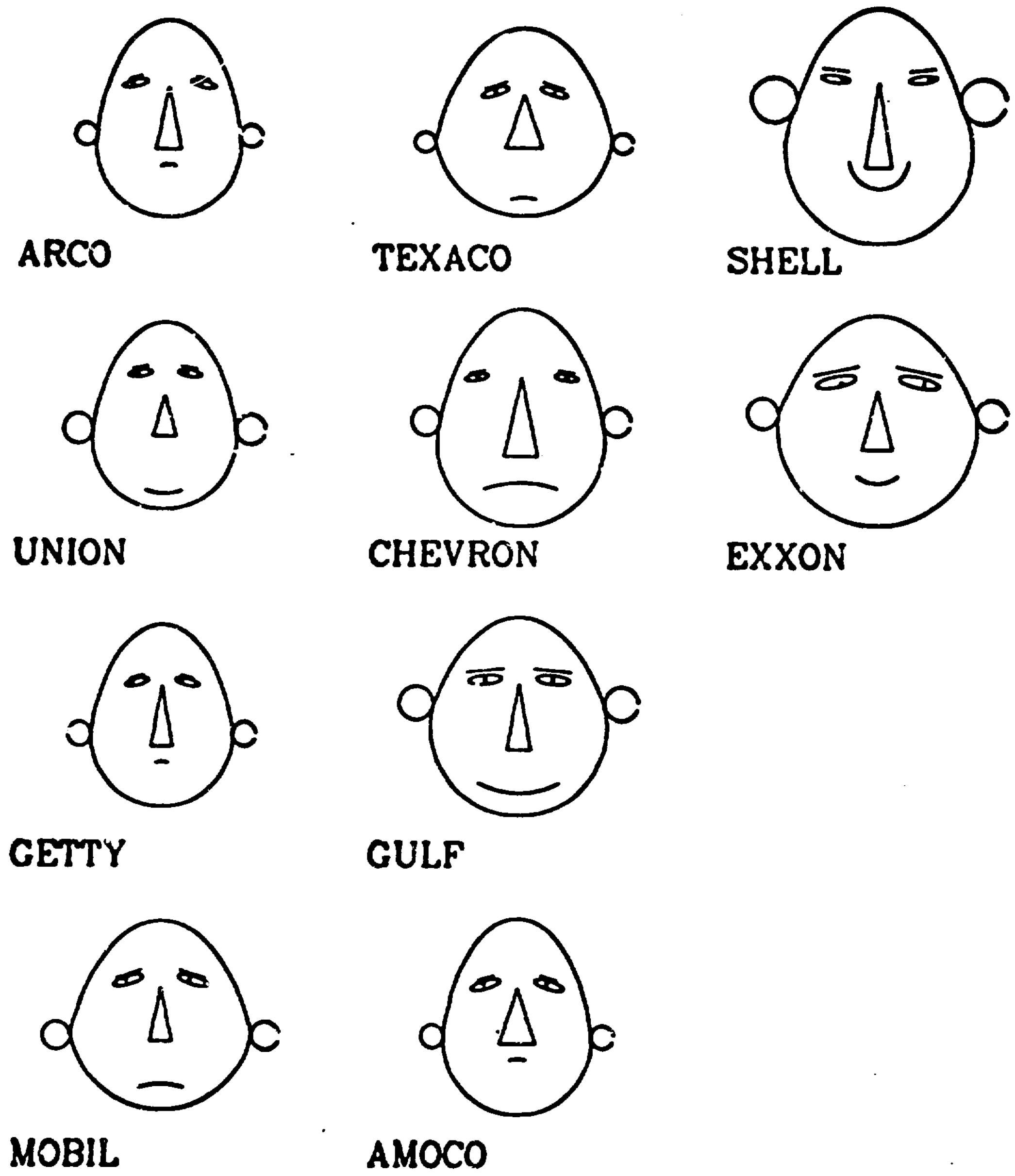


Figure 4. Chernoff Faces for 10 major oil company groups

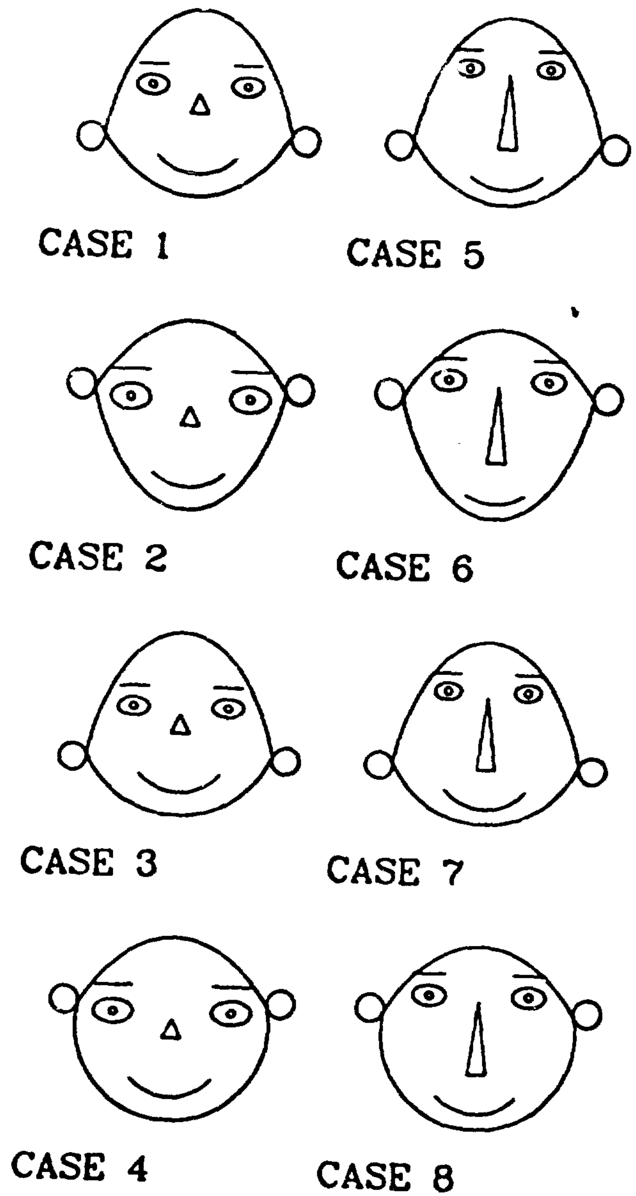


Figure 7. Facial sensitivity to nose length, lower face eccentricity and ear level

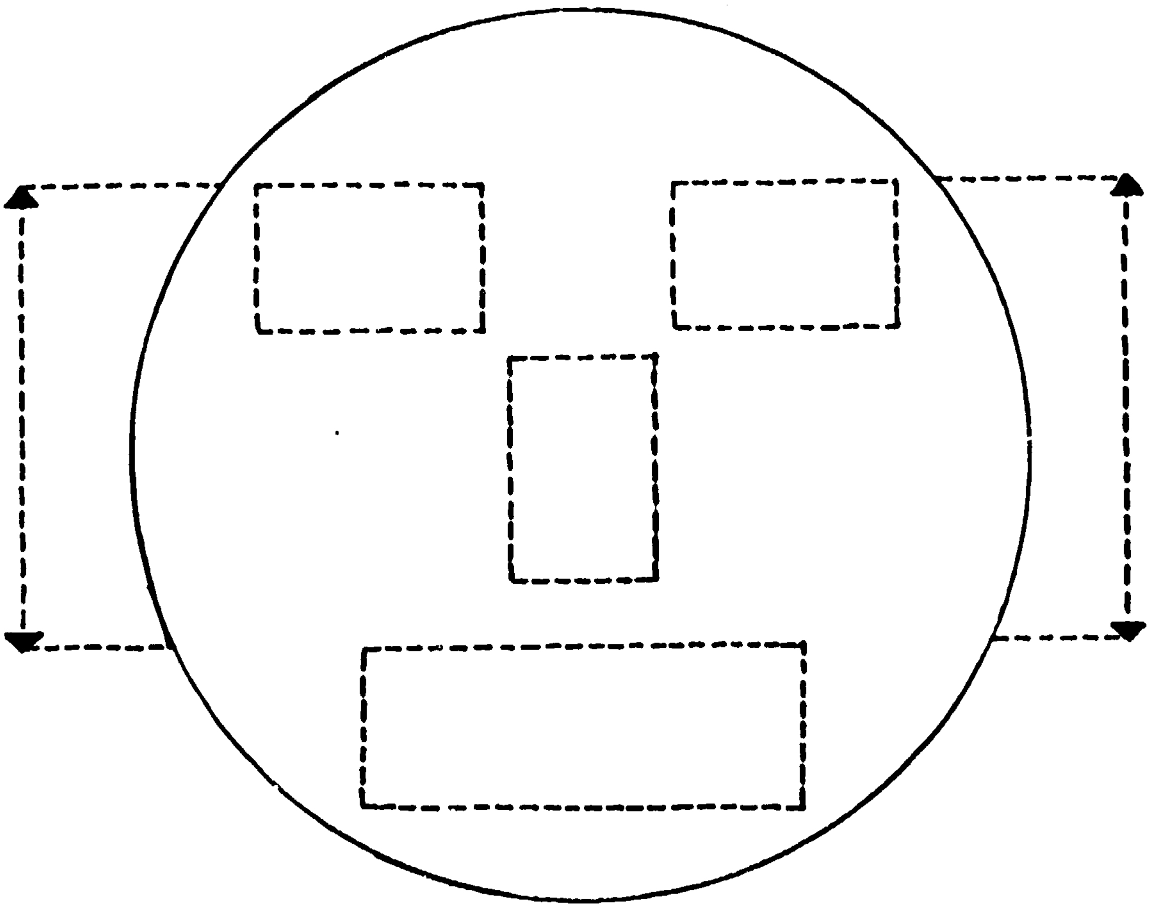


Figure 8. Suggested regions to remove dependencies